

Strategies for btagging calibration using data at CMS

The CMS Collaboration is studying several algorithms to discriminate jets coming from the hadronization of b quarks from the lighter background. These will be used to identify top quarks and in searches of the Higgs boson and non-Standard Model processes. A reliable estimate of the performance of these algorithms is therefore crucial, and methods to estimate efficiencies and mistag rates directly on data are needed. While on simulated data it was shown the btagging algorithms are shown to reach adequate performance for standard model and beyond analyses, when searching for the better b efficiency / light rejection, it is definitely not trivial to extract and validate these figures on real data. The CMS Monte Carlo simulation, even if tuned for more than 10 years, is in fact not expected to be reliable on the first data, and large discrepancies can also come from the experimental inputs on the production of heavy flavours. The CMS btagging group has prepared several strategies to extract efficiencies and rejection rates from data, which should work even on the first data (10 pb⁻¹). Three methods are currently studied. The first extracts rejection rates from light quarks looking at tracks with negative impact parameter, and using these distributions to model the mistag rate due to detector effects like resolutions, badly reconstructed tracks etc. The second method uses samples with reconstructed muons; by applying cuts on b-tagging and on the p_T of the muon relative to the jet, a system of equations can be constructed which leads to the direct extraction of efficiencies and rejections. The third method uses ttbar events with semi leptonic or fully leptonic W decays, and uses likelihood-based and event counting based methods to estimate the efficiencies. The strategies here described are studied taking into account the possible startup scenarios of LHC, and are currently being expanded to take into account the miscalibration scenarios (alignment etc) which CMS can show at the startup.

Primary author: WYSLOUCH, Boleslaw (MIT)

Presenter: KELLER, Jason (University of Nebraska-Lincoln)

Track Classification: Poster session